

Mean-field models for non-Markovian epidemics on networks: from edge-based compartmental to pairwise models

N. Sherborne¹, J.C. Miller², K.B. Blyuss^{1*}, I.Z. Kiss¹

¹ Department of Mathematics, School of Mathematical and Physical Sciences, University of Sussex, Falmer, Brighton BN1 9QH, UK

² School of Mathematics, School of Biology, and MAXIMA, Monash University, Melbourne, VIC Australia and Institute for Disease Modeling, Bellevue, Washington 98005, USA

November 15, 2016

Abstract

This paper presents a novel extension of the edge-based compartmental model for epidemics with arbitrary distributions of transmission and recovery times. Using the message passing approach we also derive a new pairwise-like model for epidemics with Markovian transmission and an arbitrary recovery period. The new pairwise-like model allows one to formally prove that the message passing and edge-based compartmental models are equivalent in the case of Markovian transmission and arbitrary recovery processes. The edge-based and message passing models are conjectured to also be equivalent for arbitrary transmission processes; we show the first step of a full proof of this. The new pairwise-like model encompasses many existing well-known models that can be obtained by appropriate reductions. It is also amenable to a relatively straightforward numerical implementation. We test the theoretical results by comparing the numerical solutions of the various pairwise-like models to results based on explicit stochastic network simulations.

1 Introduction

The use of mathematical tools to study and understand the spread of infectious diseases is an established and fruitful area of research. In their 1927 paper [1] Kermack and McKendrick established the susceptible-infected-recovered (SIR) framework which forms the basis of many models to this day. However, these early models do not consider realistic human behaviour and interactions. A particular challenge for the construction of more realistic models lies in capturing these contact patterns. Typically, in large populations each individual only interacts with a few others, and these connections determine the possible routes of disease transmission.

Moreover, studies have found significant heterogeneity in the number of contacts a single individual may have [2], as well as a tendency for two connected nodes to share further common neighbours. This clustering of individuals introduces multiple potential routes of transmission between any two individuals and provides a challenge to modellers [3, 4].

A major development which takes into account these factors has been the use of graphs or networks which provide a detailed description of these contact patterns, as opposed to assuming that each

*Corresponding author: K.Blyuss@sussex.ac.uk

individual interacts directly with every other member of the population. In network-based models individuals are represented by nodes in the network, with edges (or links) encoding the interactions between the nodes.

The resulting models of epidemics on such networks lead to a continuous time Markov chain on a very large but discrete state space. The direct analysis of such stochastic models for realistic networks is out of reach, unless strong simplifying assumptions about the networks are made. However, alternatives to deal with such models do exist. For example, it is possible to derive appropriate mean-field models that are capable of capturing the average behaviour of a stochastic model with a high degree of accuracy. These then allow one to analytically determine, explicitly or implicitly, quantities, such as the epidemic threshold or final epidemic size, and reveal how these depend on the characteristics of the network.

Deriving such deterministic mean-field models can be done in several different ways, depending on how the averaging is done. For example, averaging over all nodes and links in all possible states leads to pairwise models [5, 6], while considering each individual separately and all possible ways in which it can become infected by its neighbours leads to the message passing formalism [7]. Furthermore, choosing to average over all possible star-like structures that are typically defined by a node and all its neighbours, and also taking into account their disease status, yields the so-called effective-degree models [8]. One of the most compact mean-field models is the edge-based compartmental model which is based on considering a randomly chosen test node and working out its probability of staying susceptible, this probability is equivalent to the proportion of nodes that are susceptible in the entire population [9], see [10, 11, 12] for reviews. Of course, all of these models start from the same exact stochastic model but choose to average over a different scale, thus it is not surprising that some of these models are, in fact, equivalent [6, 13, 14, 12], as we will demonstrate later on.

Throughout this paper we restrict ourselves to unweighted, bi-directional and static networks constructed according to the configuration model (CM) [15]. Every node u is assigned a number of neighbours, known as its degree, according to a probability distribution p_k known as the *degree distribution*. This describes the probability of a randomly chosen node having degree k . One can also define the following generating functions which will be used throughout the paper

$$\begin{aligned} G_0(x) &:= \sum_k p_k x^k, & G_1(x) &:= \frac{1}{\langle k \rangle} \sum_k p_k k x^{k-1}, \\ G_2(x) &:= G_1'(x) = \frac{1}{\langle k \rangle} \sum_k p_k k(k-1) x^{k-2}, \end{aligned} \tag{1}$$

where $\langle k \rangle = G_0'(1)$ is the mean degree. $G_1(x)$ is the generating function for the *excess degree distribution*, since $k p_k / \langle k \rangle$ describes the probability that a node reached by traversing a randomly selected edge has $(k-1)$ other contacts [16]. On a tree network with no short loops the number of neighbours that an infected node can reach is determined by the excess degree of the nodes; the mean excess degree is given by $G_1'(1)$. The moment generating function $G_2(x)$ is used to trace the route of infection.

In addition to describing the underlying contact pattern, we must consider the disease characteristics. These are the force of infection or transmission and the typical length of time between an individual becoming infected and later recovering to (permanent) immunity (henceforth referred to as the recovery time). Many models assume that the underlying stochastic transmission and recovery processes are memoryless [17, 18, 6]. These assumptions lead to models which are mathematically tractable and relatively simpler to analyse when compared to models where the time to infection or the recovery time are chosen from arbitrary distributions other than the exponential.

However, when compared to data, these assumptions are often violated, and diseases can exhibit unique and non-Markovian behaviour in terms of the strength and duration of infection. For example,

recovery time distributions are usually better approximated by some peaked distribution with a well defined mean, see e.g. [19, 20, 21] and references therein. Some work has been done recently on modelling network epidemics with non-exponential distribution of infectious periods [22, 23, 24]. In this paper we will take the most general view possible by letting $\tau(a)$ and $q(a)$ denote the general probability density functions of the time to infection across a link between an infectious and susceptible node, and the recovery time of an infected/infectious node, respectively. Here, a denotes the time since the node became infected, also known as the age of infection. Once a susceptible node has been exposed to a transmission event it becomes infected immediately - a latent period can be included by setting the transmission distribution $\tau(a)$ to be zero for some initial period of time. Recovery from the disease grants lifetime immunity. Using these distributions assumes a homogeneous response to disease; whilst this restriction is not always necessary (see e.g. [25]), it is a common simplification in order to obtain a concise model.

Considering locally tree-like networks with arbitrary heterogeneous degree distributions, and a stochastic epidemic model where both the transmission and recovery processes are as general as possible, we set out to derive the most general edge-based and pairwise-like models, and to establish different model equivalences and relationships between these and existing models.

The rest of the paper is organised as follows: in the following section we describe how the message passing (MP) approach works [7] and how the resulting model is constructed. We also present a novel extension of the edge-based compartmental model (EBCM) [9] to networks with general degree distributions and for SIR epidemics with arbitrary transmission and recovery processes. In Section 3, starting from the message passing model and considering Markovian transmission, we show how one can derive a low-dimensional or compact pairwise-like model (PLM) which is more amenable to implement numerically but is still capable of modelling arbitrary recovery time and degree distributions. In Section 4 we use the new PLM to show equivalence between the MP model and EBCM for Markovian transmission, and thus to conjecture that the MP model and EBCM are equivalent in a general setting. The PLM is also reducible to many existing models for specific choices of the recovery time or degree distribution, as will be shown in Section 5. In Section 6 we compare the numerical solutions of the mean-field models to averaged results from explicit stochastic network simulations. The good agreement between these confirm and support earlier theoretical findings. The paper concludes with a brief discussion of the major results and possible future work.

2 Model summary

2.1 The message passing (MP) method

In their 2010 paper [7] Karrer and Newman introduced the message passing approach to model SIR dynamics on networks. Here, we briefly present the ideas behind their model and its assumptions. Recalling $\tau(a)$ and $q(a)$ as the densities for transmission and recovery time one can introduce a new function $f(a)$

$$f(a) = \tau(a) \int_a^\infty q(x) dx, \quad (2)$$

such that the probability that an infected node attempts to transmit the disease to a given neighbour before time t is $\int_0^t f(a) da$, since a neighbour can only transmit the disease if it has not yet recovered. Note that the result of integration of (2) over all time is equal to the overall probability of the disease being transmitted across a given network edge, commonly known as the *transmissibility*. This is an important quantity which is used in percolation models to determine the epidemic threshold and expected final epidemic size for a major outbreak [16, 26].

In order to model the dynamics of disease spread consider a test node u . This node is placed into a cavity state where it can become infected but is not able to transmit the disease to any of its neighbours. This has no effect on the probability of the node being in any given state [9]. Now consider a node v which is a neighbour of u ; the *message* is the probability that v has *not* transmitted the disease to node u by time t , denoted $H^{u \leftarrow v}(t)$. This probability is comprised of two distinct possibilities; the first possibility is that v makes no attempt to transmit the disease before t regardless of its own age of infection, given by $1 - \int_0^t f(a) da$. Alternatively, it could be that v will transmit to u at age $a < t$, but v itself was infected at some time $t_1 > t - a$ and has, therefore, not yet attempted to transmit the disease to its neighbour u . This requires v to have initially been susceptible (with probability z) and to have escaped transmission from each of its neighbours (excluding u) until at least time $(t - a)$, which is exactly $z \int_0^t f(a) \prod_{w \in \mathcal{N}(v) \setminus u} H^{v \leftarrow w}(t - a) da$, where $\mathcal{N}(v)$ denotes the set of neighbours of v . Hence, combining these two gives

$$H^{u \leftarrow v}(t) = 1 - \int_0^t f(a) \left[1 - z \prod_{w \in \mathcal{N}(v) \setminus u} H^{v \leftarrow w}(t - a) \right] da. \quad (3)$$

In principle, one could calculate (3) for all edges (in both directions) to find a full solution for the proportion of the population that is susceptible, infected or removed at any time t . For example, the probability u is susceptible is the product of $H^{u \leftarrow w}(t)$ across all neighbours $w \in \mathcal{N}(u)$ multiplied by the probability that it was initially susceptible, z . On a single fixed finite tree network, solving (3) for all edges will, in fact, yield the exact solution of the stochastic epidemic [7]. The size of such a system of equations would be twice the number of all edges in the network (since both $H^{u \leftarrow v}(t)$ and $H^{v \leftarrow u}(t)$ would need to be calculated).

However, for CM networks it is possible to modify (3) so that it performs as an averaged probability of all edges across the whole network, this is denoted $H_1(t)$ [7]. In order to account for degree heterogeneity, the product in (3) is replaced by the excess degree distribution $G_1(H_1(t - a))$. Thus, we now have

$$H_1(t) = 1 - \int_0^t f(a) [1 - z G_1(H_1(t - a))] da, \quad (4)$$

with $H_1(0) = 1$. Now the somewhat difficult part is to obtain and solve, at least numerically, the differential or integro-differential equation for $H_1(t)$. For the purely Markovian case, with transmission and recovery parameters β and γ respectively, this is

$$\frac{dH_1}{dt} = \gamma - (\beta + \gamma)H_1(t) + \beta z G_1(H_1(t)),$$

where z is the fraction of the population which was initially susceptible at time $t = 0$ [7]. However, the precise form of this equation is not universal, it depends on the particular choice of the transmission and recovery processes. The proportion of susceptible, infected and recovered individuals at any time t are then given, in terms of the message $H_1(t)$, as

$$\begin{aligned} \langle S \rangle(t) &= z G_0(H_1(t)), \\ \langle R \rangle(t) &= \int_0^t q(a) [1 - \langle S \rangle(t - a)] da, \\ \langle I \rangle(t) &= 1 - \langle S \rangle(t) - \langle R \rangle(t). \end{aligned} \quad (5)$$

For CM networks, as the size of the network increases to infinity, the length of the shortest loops also diverges to infinity, and, therefore, the network becomes locally tree-like. The result of this is

that an MP model with the average message H_1 is exact when the stochastic epidemic is considered on the ensemble of CM networks [7]. Although this approach is theoretically able to model dynamics for arbitrary choices of transmission and recovery processes, the need to find a numerically solvable differential equation for H_1 in (4) has restricted the use of MP, and numerical examples are limited, see [7, 25] for several examples where output from the MP model is compared to results based on simulations.

2.2 EBCM for general transmission and recovery processes

The edge-based compartmental model has been used for Markovian dynamics [9]. We introduce a new extended EBCM which generalises the method to general transmission and recovery processes.

The EBCM uses the instantaneous rates of transmission and recovery given by the *hazard functions* rather than the raw densities $\tau(a)$ and $q(a)$. These are defined as

$$\zeta(a) := \frac{\tau(a)}{\xi_\tau(a)}, \quad \text{and} \quad \rho(a) := \frac{q(a)}{\xi_q(a)}, \quad (6)$$

where $\xi_\tau(a)$ and $\xi_q(a)$ are the respective *survival functions*

$$\xi_\tau(a) = \int_a^\infty \tau(\hat{a}) d\hat{a}, \quad \text{and} \quad \xi_q(a) = \int_a^\infty q(\hat{a}) d\hat{a}. \quad (7)$$

It is possible to express these survival functions in an alternative form which will be useful later. These are given in the following lemma.

Lemma 1. *Given $\xi_\tau(a)$ and $\xi_q(a)$ as defined in (7), and the relations in (6), one can write*

$$\xi_\tau(a) = e^{-\int_0^a \zeta(\hat{a}) d\hat{a}} \quad \text{and} \quad \xi_q(a) = e^{-\int_0^a \rho(\hat{a}) d\hat{a}}. \quad (8)$$

Proof. From (6) and (7) we see that

$$\xi_\tau(a) = \int_a^\infty \zeta(\hat{a}) \xi_\tau(\hat{a}) d\hat{a}.$$

Differentiating this equation with respect to a yields a first order differential equation

$$\frac{d\xi_\tau}{da} = -\zeta(a)\xi_\tau(a),$$

whose solution with the initial condition $\xi_\tau(0) = 1$ gives

$$\xi_\tau(a) = e^{-\int_0^a \zeta(\hat{a}) d\hat{a}}.$$

Applying the same steps to $\xi_q(a)$ completes the proof. ■

All of these disease variables and related functions are summarised in Table 1. As before, the contact network is a CM network with degree distribution and generating functions as defined in (1). The basis of the EBCM revolves around finding the probability that a random test node is in a susceptible, infected or recovered state at time t ; as this test node is chosen at random these probabilities are equal to the proportions of the population in each state at time t , denoted $S(t)$,

Variable	Definition
$\tau(a)$	The density of the transmission process.
$q(a)$	The density of the recovery process.
$\xi_\tau(a)$	The <i>survival function</i> of the transmission process. The probability that an infected node of age a has not yet attempted to transmit the disease along a given edge: $\int_a^\infty \tau(x) dx$.
$\xi_q(a)$	The <i>survival function</i> of the recovery process. The probability that an infected node reaches at least age a before recovering: $\int_a^\infty q(x) dx$.
$\zeta(a)$	The <i>hazard function</i> of the transmission process. The probability of an edge of age a transmitting in a small interval of time $(a, a + \Delta a)$: $\frac{\tau(a)}{\xi_\tau(a)}$.
$\rho(a)$	The <i>hazard function</i> of the recovery process. The probability of an infected node of age a recovering in a small interval of time $(a, a + \Delta a)$: $\frac{q(a)}{\xi_q(a)}$.
$f(a)$	The probability that, in a small interval, an infectious contact is made by an infected node of age a : $\tau(a) \int_a^\infty q(x) dx$.
$g(a)$	The probability that, in a small interval, an infectious node of age a recovers, without attempting to transmit the disease to a given neighbour: $q(a) \int_a^\infty \tau(x) dx$.

Table 1: The variables and functions describing the transmission and recovery processes.

$I(t)$ and $R(t)$ respectively. Again, the test node is placed into a cavity state, and the probability of remaining susceptible until time t is the probability of the test node being initially susceptible and escaping transmission from each of its neighbours up to time t . This concept is similar to the notion and use of H_1 in MP models. Recovery is modelled using age-structured differential equations.

The first important quantity is $\Theta(t)$, defined in a manner similar to $H_1(t)$ in (4) as the probability that the representative test node has not received transmission from a given neighbour by time t . This approach then differs from MP by directly expressing a differential equation for the dynamics of Θ . The model is known as “edge-based” because it considers the state of the neighbours of the test node; the densities $\Phi_S(t)$, $\Phi_I(t)$ and $\Phi_R(t)$ describe the probability that at time t a random neighbour of the test node is (i) still susceptible, (ii) infected but has not attempted to transmit the disease to the test node, (iii) recovered, and it did not transmit to the test node whilst it was infected. Since the age of infection is crucial in determining the hazard rates we also introduce the variables $i(t, a)$ and $\phi_i(t, a)$ as the densities of infected nodes with the age of infection a , and infected neighbours who have not transmitted to the test node and have age a , respectively. Thus, it is clear that $I(t) = \int_0^t i(t, a) da$ and $\Phi_I(t) = \int_0^t \phi_I(t, a) da$. These variables are summarised in Table 2. The new EBCM is presented below for the case where the disease is introduced by randomly infecting a chosen fraction $1 - z$ of the

Variable	Definition
$\Theta(t)$	The probability an initially susceptible test node has not received a transmission from a random neighbour by time t .
$\Phi_S(t)$	The probability that a random neighbour of a test node u is still susceptible.
$\Phi_I(t)$	The probability that a random neighbour of a test node u is infected, but has not transmitted to u .
$\phi_I(t, a)$	The density for a random neighbour of a test node u to be infected, have not transmitted to u by time t and be a units of time into its infection, $\Phi_I(t) = \int_0^t \phi_I(t, a) da$.
$\Phi_R(t)$	The probability a random neighbour of a test node u has been infected and recovered without transmitting to u .
$S(t)$	The density of susceptible nodes.
$I(t)$	The density of infected nodes.
$i(t, a)$	The density of infected nodes that were infected at time $(t - a)$.
$R(t)$	The density of recovered nodes.
$G_1(x)$	The probability generating function of the excess degree distribution: $\frac{1}{\langle k \rangle} \sum_{k=0}^{\infty} p_k k x^{(k-1)}$.
$G_2(x)$	The derivative of the probability generating function of the excess degree distribution: $\frac{1}{\langle k \rangle} \sum_{k=0}^{\infty} p_k k(k-1)x^{(k-2)}$.

Table 2: The list of variables in the EBCM.

population at time 0.

$$\Phi_S(t) = zG_1(\Theta(t)), \quad (9a)$$

$$\begin{aligned} \phi_I(t, 0) &= -\dot{\Phi}_S(t), \\ &= (1 - z)\delta(t) + zG_2(\Theta(t)) \int_0^t \zeta(a)\phi_I(t, a) da, \end{aligned} \quad (9b)$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) \phi_I(t, a) = -[\zeta(a) + \rho(a)] \phi_I(t, a), \quad 0 < a \leq t, \quad (9c)$$

$$\Phi_I(t) = \int_0^t \phi_I(t, a) da, \quad (9d)$$

$$\frac{d\Theta(t)}{dt} = - \int_0^t \zeta(a)\phi_I(t, a) da, \quad (9e)$$

$$\Phi_R(t) = \Theta - \Phi_S - \Phi_I, \quad (9f)$$

$$S(t) = zG_0(\Theta(t)), \quad (9g)$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) i(t, a) = -\rho(a)i(t, a), \quad 0 < a \leq t, \quad (9h)$$

$$\begin{aligned} i(t, 0) &= -\dot{S}(t), \\ &= (1 - z)\delta(t) + \langle k \rangle zG_1(\Theta(t)) \int_0^t \zeta(a)\phi_I(t, a) da, \end{aligned} \quad (9i)$$

$$I(t) = \int_0^t i(t, a) da, \quad (9j)$$

$$R(t) = 1 - S - I. \quad (9k)$$

Where $\delta(t)$ is the Dirac δ -function, and we have the initial condition

$$\Theta(0) = 1,$$

and where $\langle k \rangle$ denotes the average node degree.

Compared to the MP method, the new EBCM is more universal in the sense that different choices of $\tau(a)$ and $q(a)$ can be substituted straight into the general model rather than having to consider how the equation for H_1 changes as a result of the choices being made. This arguably makes the resultant EBCM more amenable to numerical implementation, especially in cases where the evolution equation for $H_1(t)$ is difficult to obtain.

As it turns out, setting up the model in such a general way allows one to derive an implicit analytical relation for the final epidemic size, as presented in the following result.

Theorem 1. *The final size of the epidemic $r_\infty = R(\infty)$ in the EBCM (9) with a vanishingly small proportion of infected nodes at time $t = 0$ is*

$$r_\infty = 1 - G_0(\Theta_\infty),$$

where Θ_∞ solves the equation

$$\Theta_\infty = 1 - \tilde{T} + \tilde{T}G_1(\Theta_\infty),$$

and $\tilde{T} = \int_0^\infty \zeta(a) \exp(-\int_0^a [\zeta(\hat{a}) + \rho(\hat{a})] d\hat{a}) da$ is known as the transmissibility of the disease. It is the probability that the disease is transmitted along an edge (in isolation) before recovery.

Proof. From (9g), for $z \rightarrow 1$ and $t \rightarrow \infty$ it immediately follows that

$$r_\infty = 1 - S(\infty) = 1 - G_0(\Theta(\infty)).$$

Furthermore, we have

$$\Theta(\infty) = 1 - \int_0^\infty \int_0^t \zeta(a) \phi_I(t, a) da dt.$$

Interchanging the order of integration yields

$$\Theta(\infty) = 1 - \int_0^\infty \int_a^\infty \zeta(a) \phi_I(t, a) dt da.$$

Setting $u = t - a$ and noting that $\phi_I(t, a) = \phi_I(u, 0) \exp(-\int_0^a [\zeta(\hat{a}) + \rho(\hat{a})] d\hat{a})$ yields

$$\begin{aligned} \Theta(\infty) &= 1 - \int_0^\infty \int_0^\infty \phi_I(u, 0) \zeta(a) e^{-\int_0^a [\zeta(\hat{a}) + \rho(\hat{a})] d\hat{a}} du da \\ &= 1 - \left[\int_0^\infty \phi_I(u, 0) du \right] \int_0^\infty \zeta(a) e^{-\int_0^a [\zeta(\hat{a}) + \rho(\hat{a})] d\hat{a}} da \\ &= 1 + \left[\int_0^\infty \dot{\Phi}_S(u) du \right] \int_0^\infty \zeta(a) e^{-\int_0^a [\zeta(\hat{a}) + \rho(\hat{a})] d\hat{a}} da \\ &= 1 + (\Phi_S(\infty) - \Phi_S(0)) \tilde{T} \\ &= 1 + \tilde{T} G_1(\Theta(\infty)) - \tilde{T}. \end{aligned}$$

■

The result in Theorem 1 corresponds to well-known results based on tools from percolation theory [16, 26, 27], and it is equivalent to the final epidemic size obtained for MP models [7]. This provides good evidence that the new EBCM (9) is an accurate representation of the true expected dynamics of an SIR epidemic starting from a vanishingly small number of initially infected seeds.

3 Derivation of the new pairwise-like model

The number of individuals becoming infected is related to the number of edges connecting susceptible nodes to infected neighbours. Pairwise models traditionally construct differential equations for the expected numbers of such edges, which themselves depend on the numbers of triples in certain states (e.g. susceptible-susceptible-infected). To break this dependence a moment closure approximation is commonly used to express the number of triples in terms of pairs and individuals [5]. Recently, Wilkinson and Sharkey [25] and Wilkinson et al. [28] have shown that for regular tree networks exact pairwise models can be derived from the MP model when the transmission process is assumed to be Markovian. Here we use similar methods with the notation from Section 2.1 to extend this result to heterogeneous networks.

Firstly, we define the new variable $\langle SI \rangle(t)$ as the proportion of edges in the network which connect a susceptible node to an infected one at time t . This can be defined in terms of existing quantities; first the susceptible node must have been initially susceptible and escaped infection from all other neighbours, given by $zG_1(H_1(t))$. This must be multiplied by the probability that the remaining neighbour of the susceptible node is infected and has not yet transmitted the disease to this neighbour. To find this probability it is easier to calculate all other possibilities and subtract them from one. These possibilities are: (i) the neighbour is still susceptible, (ii) the neighbour has already transmitted the

disease, (iii) the neighbour was infected but recovered without transmitting the disease. Combining these gives

$$\begin{aligned} \langle SI \rangle(t) &= zG_1(H_1(t)) \left[1 - zG_1(H_1(t)) \right. \\ &\quad \left. - \int_0^t \{f(a) + g(a)\} [1 - zG_1(H_1(t-a))] da \right], \end{aligned} \quad (10a)$$

$$\begin{aligned} &= zG_1(H_1(t)) \left[H_1(t) - zG_1(H_1(t)) \right. \\ &\quad \left. - \int_0^t g(a) [1 - zG_1(H_1(t-a))] da \right], \end{aligned} \quad (10b)$$

$$(10c)$$

where

$$g(a) := q(a) \int_a^\infty \tau(x) dx, \quad (11)$$

is the probability of an infected node recovering in the interval $(a, a + \Delta a)$ without transmitting to a given neighbour. The corresponding population-level quantity is given by

$$[SI](t) = \langle k \rangle N \langle SI \rangle(t), \quad (12)$$

where N denotes the total size of the population.

We now assume that the time to transmission is exponentially distributed, i.e. $\tau(a) = \beta e^{-\beta a}$. Substituting this into (4), changing the variable to $t' = t - a$, and using the Leibniz rule gives

$$\begin{aligned} \frac{dH_1}{dt} &= -\beta \left[1 - zG_1(H_1(t)) - \int_0^t q(t-t') e^{-\beta(t-t')} [1 - zG_1(H_1(t'))] da \right. \\ &\quad \left. - \int_0^t \beta e^{-\beta(t-t')} \left(\int_{t-t'}^\infty q(x) dx \right) [1 - zG_1(H_1(t'))] dt' \right] \\ &= -\beta \left[1 - zG_1(H_1(t)) - \int_0^t \{f(a) + g(a)\} [1 - zG_1(H_1(t'))] da \right] \\ &= -\beta \frac{\langle SI \rangle(t)}{zG_1(H_1(t))} \end{aligned} \quad (13)$$

$$= -\beta \frac{[SI]}{z \langle k \rangle N G_1(H_1)}. \quad (14)$$

For the infected population, using (5) and identities, such as $[S](t) = N \langle S \rangle(t)$, leads to

$$[\dot{I}] = -[\dot{S}] - [\dot{R}] \quad (15)$$

$$= \beta[SI] - \beta \int_0^t q(a)[SI](t-a) da - q(t)N(1-z). \quad (16)$$

The majority of pairwise epidemic models retain an explicit differential equation for the prevalence [6, 28]. However, we choose to integrate (16) to reduce the number of differential equations which must be integrated numerically. By noting that $[SI] = \beta \int_0^t [\dot{SI}](t-a) \xi_q(a) da$ and $q(a) = -\xi_q'(a)$ we have

$$[\dot{I}] = \beta \int_0^t \left([\dot{SI}](t-a) \xi_q(a) + \xi_q'(a)[SI](t-a) \right) da - q(t)N(1-z),$$

which is the result of differentiating

$$[I] = \beta \int_0^t [SI](t-a)\xi_q(a) da + N(1-z)\xi_q(t). \quad (17)$$

Whilst (16) facilitates easier comparison to existing models we will use its equivalent representation (17) for computational efficiency.

For the variable $\langle SI \rangle$ the calculation is more laborious; working term-by-term from (10b) and using the new relation (14) one obtains

$$\begin{aligned} \langle \dot{SI} \rangle = & -zG_2(H_1) \left(\beta \frac{[SI]}{z\langle k \rangle N G_1(H_1(t))} \right) \left[\cdot \right] \\ & - \beta \frac{[SI]}{\langle k \rangle N} + z\beta[SI] \frac{G_2(H_1(t))}{\langle k \rangle N} - q(t)e^{-\beta t}(1-z)zG_1(H_1(t)) \\ & - z\beta \int_0^t q(a)e^{-\beta a}[SI](t-a)G_2(H_1(t-a)) \frac{G_1(H_1(t))}{N\langle k \rangle G_1(H_1(t-a))} da, \end{aligned}$$

where $[\cdot]$ denotes the large bracket in (10b), and the Leibniz rule has been used again to resolve the integral term. Finally, based on (10), $[\cdot] = \frac{\langle SI \rangle}{zG_1(H_1)}$, which allows us to eliminate $[\cdot]$ and replace it with a term involving $\langle SI \rangle$. Then, multiplying through by $\langle k \rangle N$ one obtains $[SI]$ in (18) below.

$$\dot{H}_1 = -\beta \frac{[SI]}{z\langle k \rangle N G_1(H_1)}, \quad (18a)$$

$$\begin{aligned} [SI] = & -\beta[SI][SI] \frac{G_2(H_1)}{z\langle k \rangle N [G_1(H_1)]^2} - \beta[SI] \\ & + z\beta[SI]G_2(H_1) - q(t)e^{-\beta t}(1-z)zG_1(H_1)\langle k \rangle N \\ & - z\beta \int_0^t q(a)e^{-\beta a}[SI](t-a)G_2(H_1(t-a)) \frac{G_1(H_1(t))}{G_1(H_1(t-a))} da, \end{aligned} \quad (18b)$$

$$[I] = \beta \int_0^t [SI](t-a)\xi_q(a) da + N(1-z)\xi_q(t). \quad (18c)$$

At any time t the expected number of susceptibles can be found as $[S](t) = zNG_0(H_1(t))$. We call the system (18) the pairwise-like model (PLM). The major benefits of this model are its compact size, requiring only two differential equations to be solved, and the fact that retaining the concept of the “message” from the MP model has meant that moment closure approximations are not necessary. The PLM is a concise, flexible and exact model for SIR dynamics on CM networks. Moreover, the numerical solution of such an ODE system is straightforward to compute.

There are elements of (18) that are similar to recent models [6, 28, 29], which helps us identify links between different methodologies and promote a greater and more unified understanding of this area. These connections are explored in detail in Section 5.

4 Model equivalence

Before exploring how the new PLM (18) reduces to recent and classical models, we use it to show equivalence between the MP model and the EBCM introduced in Section 2 in the special case of Markovian transmission but an arbitrary recovery process. Since the PLM is derived directly from

the MP model, it is only necessary to show equivalence between the PLM and EBCM. Note that whilst all variables in both the MP model and EBCM are based on proportions, the PLM deals with population-level quantities, and hence we will show that the corresponding quantities (e.g. proportion and population of infection) share identical dynamics and differ only by constants, such as the total size of the population, N . This equivalence is summarised in the following theorem.

Theorem 2. *Given an exponentially distributed transmission process such that $\tau(a) = \beta e^{-\beta a}$ with some $\beta > 0$, the dynamics of the EBCM (9) are equivalent to that of the PLM and, therefore, are identical to the MP model and stochastically exact on CM networks as the population size tends to infinity.*

Proof. The PLM (18) has been derived directly from the MP model (5), so to prove equivalence between the MP and EBCM systems it will suffice to show that the PLM can be independently derived from the EBCM (9). We introduce the new variables

$$\{SI\} = zN\langle k \rangle G_1(\Theta(t))\Phi_I(t), \quad (19a)$$

$$\{I\} = NI, \quad (19b)$$

We then derive equations for $\{SI\}$ and $\{I\}$ and show that

$$\frac{d}{dt}\Theta = -\beta \frac{\{SI\}}{z\langle k \rangle N G_1(\Theta(t))}. \quad (19c)$$

The resulting system of equations will be identical to those of system (18) with Θ playing the role of H_1 , $\{SI\}$ playing the role of $[SI]$ and $\{I\}$ playing the role of $[I]$. Thus the PLM and EBCM systems are equivalent.

The relation (19c) immediately follows from $\frac{d\Theta}{dt} = -\beta\Phi_I$ and substituting for Φ_I in terms of $\{SI\}$.

To calculate $\frac{d\{SI\}}{dt}$ we first require $\frac{d}{dt}\Phi_I(t)$, from (9) we have

$$\begin{aligned} \frac{d\Phi_I(t)}{dt} &= \frac{d}{dt} \int_0^t \phi_I(t, a) da \\ &= \int_0^t \frac{\partial}{\partial t} \phi_I(t, a) da + \phi_I(t, t) \\ &= - \int_0^t \left[(\beta + \rho(a))\phi_I(t, a) + \frac{\partial}{\partial a} \phi_I(t, a) \right] da + \phi_I(t, t) \\ &= -\beta\Phi_I(t) + \phi_I(t, 0) - \int_0^t \rho(a)\phi_I(t, a) da. \end{aligned}$$

Now to derive the evolution equation for $\{SI\}$ we start by differentiating both sides of (19a):

$$\begin{aligned} \frac{d\{SI\}}{dt} &= zN\langle k \rangle \frac{d}{dt} [G_1(\Theta(t))\Phi_I(t)] \\ &= zN\langle k \rangle \left[-\beta\Phi_I(t)G_1(\Theta(t))\Phi_I(t) + G_1(\Theta(t))\frac{d}{dt}\Phi_I(t) \right] \\ &= -\beta \frac{\{SI\}\{SI\}G_1(\Theta)}{zN\langle k \rangle G_1(\Theta(t))G_1(\Theta(t))} + zN\langle k \rangle G_1(\Theta) \frac{d}{dt}\Phi_I(t) \end{aligned}$$

The first term already matches the first term of (18b). We now explore the remaining terms

$$\begin{aligned}
zN\langle k \rangle G_1(\Theta) \frac{d}{dt} \Phi_I(t) &= zN\langle k \rangle G_1(\Theta(t)) \left[-\beta \Phi_I(t) + \phi_I(t, 0) - \int_0^t \rho(a) \phi_I(t, a) da \right], \\
&= -\beta \{SI\} + zN\langle k \rangle G_1(\Theta(t)) [(1-z)\delta(t) \\
&\quad + zG_2(\Theta(t)) \int_0^t \beta \phi_I(t, a) da - \int_0^t \rho(a) \phi_I(t, a) da], \\
&= -\beta \{SI\} + z\beta \{SI\} G_2(\Theta(t)) \Phi_I(t) \\
&\quad - zN\langle k \rangle G_1(\Theta(t)) \int_0^t \rho(a) \phi_I(t, a) da.
\end{aligned}$$

All that remains is to rewrite the final term in terms of the new variables. From (9c) we can use an integrating factor of $\exp\left(\int_0^a [\zeta(\hat{a}) + \rho(\hat{a})] d\hat{a}\right)$ to find

$$\phi_I(t, a) = e^{-\beta a} \xi_q(a) \phi_I(t-a, 0), \quad (20)$$

using the relation $\xi_q(a) = \exp\left(-\int_0^a \rho(\hat{a}) d\hat{a}\right)$ found in Lemma 1. We can progress further using (9b) to give

$$\begin{aligned}
\phi_I(t, a) &= e^{-\beta a} \xi_q(a) \left((1-z)\delta(t-a) + zG_2(\Theta(t-a)) \int_0^{t-a} \beta \phi_I(t-a, \hat{a}) d\hat{a} \right) \\
&= e^{-\beta a} \xi_q(a) [(1-z)\delta(t-a) + zG_2(\Theta(t-a)) \beta \Phi_I(t-a)].
\end{aligned} \quad (21)$$

As an alternative, we offer a graphical description of (21) in Fig. 1. Recall $q(a) = \rho(a)\xi_q(a)$, we have

$$\begin{aligned}
\int_0^t \rho(a) \phi_I(t, a) da &= \int_0^t \left\{ e^{-\beta a} q(a) [(1-z)\delta(t-a) \right. \\
&\quad \left. + z\beta G_2(\Theta(t-a)) \beta \Phi_I(t-a)] \right\} da \\
&= (1-z)e^{-\beta t} q(t) \\
&\quad + z\beta \int_0^t e^{-\beta a} q(a) G_2(\Theta(t-a)) \Phi_I(t-a) da.
\end{aligned}$$

Substituting in $\Phi_I(t-a) = \{SI\}(t-a)/zN\langle k \rangle G_1(\Theta(t-a))$ gives

$$\begin{aligned}
zN\langle k \rangle G_1(\Theta(t)) \int_0^t \rho(a) \phi_I(t, a) da &= z(1-z)N\langle k \rangle G_1(\Theta(t)) e^{-\beta t} q(t) \\
&\quad + zG_1(\Theta(t)) \beta \int_0^t q(a) e^{-\beta a} \frac{G_2(\Theta(t-a)) \{SI\}(t-a)}{G_1(\Theta(t-a))} da.
\end{aligned}$$

Combining these results, the equation for $\frac{d}{dt}\{SI\}$ matches the equation for $\frac{d}{dt}[SI]$ if we replace $[SI]$ by $\{SI\}$ and H_1 by Θ .

We now turn to the equation for $\{I\}$. We will use

$$\begin{aligned}
i(t, a) &= i(t-a, 0) \xi_q(a) \\
&= (1-z)\delta(t-a) \xi_q(a) + z\langle k \rangle G_1(\Theta(t-a)) \xi_q(a) \int_0^{t-a} \beta \phi_I(t-a, \hat{a}) d\hat{a} \\
&= (1-z)\delta(t-a) \xi_q(a) + z\langle k \rangle G_1(\Theta(t-a)) \beta \Phi_I(t-a) \xi_q(a).
\end{aligned}$$

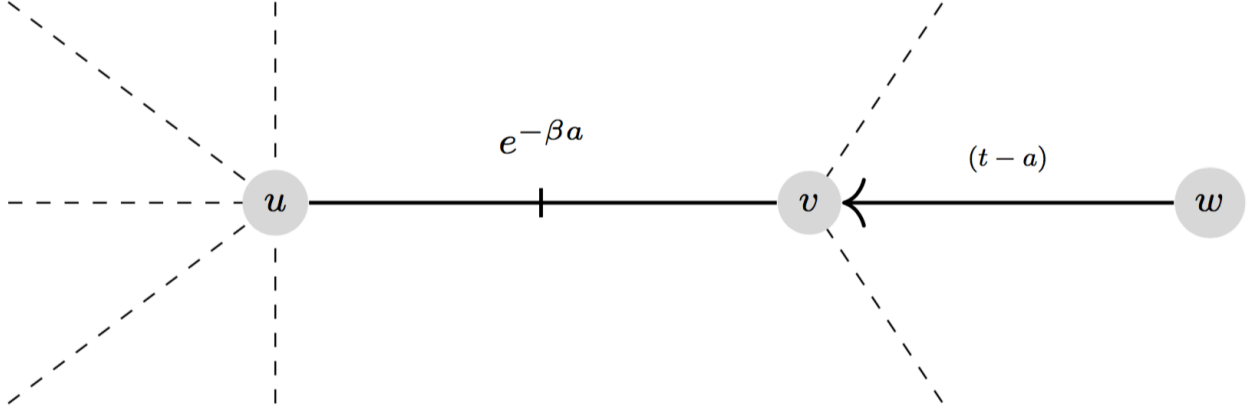


Figure 1: Consider the node labelled u as the test node and thus in a cavity state. For its link with node v to contribute to $\phi_I(t, a)$ it must be the case that v received transmission from some neighbour w at time $(t - a)$. If $t - a = 0$ then this is equal to the initial proportion of infected nodes. Otherwise we take the probability of a transmission event a time ago; which is $\beta\Phi_I(t - a)$. For v to have been successfully infected at this time it must have been susceptible until that point, since two of its neighbours will not have transmitted before this time (u is in a cavity state and w will transmit at $(t - a)$) the probability of this is $zG_2(H_1(t - a))$ for $t > a$, illustrated by the dashed lines. Regardless of how neighbour v was infected, the probability of v not transmitting to u before time t is $e^{-\beta a}$ since the transmission process is Poisson. Finally, the neighbour v must still be infected at age a , which is given by the survival function $\xi_q(a)$.

$$\begin{aligned}
\{I\} &= N \int_0^t i(t, a) da \\
&= N \int_0^t (1 - z)\delta(t - a)\xi_q(a) + z\langle k \rangle G_1(\Theta(t - a))\beta\Phi_I(t - a)\xi_q(a) da \\
&= N(1 - z)\xi_q(t) + zN\langle k \rangle\beta \int_0^t G_1(\Theta(t - a))\Phi_I(t - a)\xi_q(a) da \\
&= N(1 - z)\xi_q(t) + \beta \int_0^t \{SI\}(t - a)\xi_q(a) da.
\end{aligned}$$

Therefore, the expressions for $\{I\}$ and $[I]$ also coincide. Thus we have shown that the system of equations for $\{SI\}$, $\{I\}$, and Θ is identical to the system for $[SI]$, $[I]$, and H_1 . This completes the proof. \blacksquare

This equivalence shows that in the special case of Markovian transmission the EBCM (9) becomes exact on an ensemble of CM networks where the network size tends to infinity. Moreover, it suggests that the EBCM and MP models may indeed be equivalent for a general transmission process. The main step required to prove the general result is to show that H_1 and Θ satisfy the same evolution

equation. To do this we take

$$\begin{aligned}
\frac{dH_1}{dt} &= -f(t)[1 - zG_1(H_1(0))] - \int_0^t f(a) \left[-zG_2(H_1(t-a)) \frac{dH_1(t-a)}{dt} \right] da \\
&= -f(t)[1 - zG_1(1)] + \int_0^t f(a) \left[zG_2(H_1(t-a)) \frac{dH_1(t-a)}{dt} \right] da \\
&= -f(t)(1-z) + \int_0^t f(a) \left[zG_2(H_1(t-a)) \frac{dH_1(t-a)}{dt} \right] da.
\end{aligned} \tag{22}$$

The dynamics of Θ is governed by the following equation

$$\frac{d\Theta}{dt} = - \int_0^t \zeta(a) \phi_I(t, a) da.$$

Expression for $\phi_I(t, a)$ can be obtained from the generalised form of (20)

$$\phi_I(t, a) = \phi_I(t-a, 0) e^{-\int_0^a [\zeta(\hat{a}) + \rho(\hat{a})] d\hat{a}}.$$

Introducing $\hat{f}(a) := \zeta(a) e^{-\int_0^a [\zeta(\hat{a}) + \rho(\hat{a})] d\hat{a}}$ and using

$$\phi_I(t-a, 0) = (1-z)\delta(t-a) - zG_2(\Theta(t-a)) \frac{d\Theta(t-a)}{dt},$$

we have

$$\begin{aligned}
\frac{d\Theta}{dt} &= - \int_0^t \hat{f}(a) \left[(1-z)\delta(t-a) - zG_2(\Theta(t-a)) \frac{d\Theta(t-a)}{dt} \right] da \\
&= -\hat{f}(a)(1-z) + z \int_0^t \hat{f}(a) G_2(\Theta(t-a)) \frac{d\Theta(t-a)}{dt} da.
\end{aligned} \tag{23}$$

Thus, $H_1(t)$ and $\Theta(t)$ have the same dynamics if one can show that $f(a) = \hat{f}(a)$. From the definition of $f(a)$ and using the result of Lemma 1, we obtain

$$f(a) = \zeta(a)\xi_\tau(a)\xi_q(a) = \zeta(a) e^{-\int_0^a [\zeta(\hat{a}) + \rho(\hat{a})] d\hat{a}} = \hat{f}(a).$$

Since H_1 and Θ have the same initial condition, this implies that $H_1(t)$ and $\Theta(t)$ will exhibit identical dynamics for general transmission and recovery processes. The equivalence between the models is illustrated in the top half of Fig. 2.

5 Derivation of existing models from the PLM for regular networks and different recovery time distributions

In this section we illustrate how the new PLM (18) can be reduced to various existing models for appropriate choices of the recovery time and degree distributions. By illustrating this we emphasise that while the MP model and EBCM are probably the most competent models, the PLM is useful and plays the role of linking the most advanced to the simpler models.

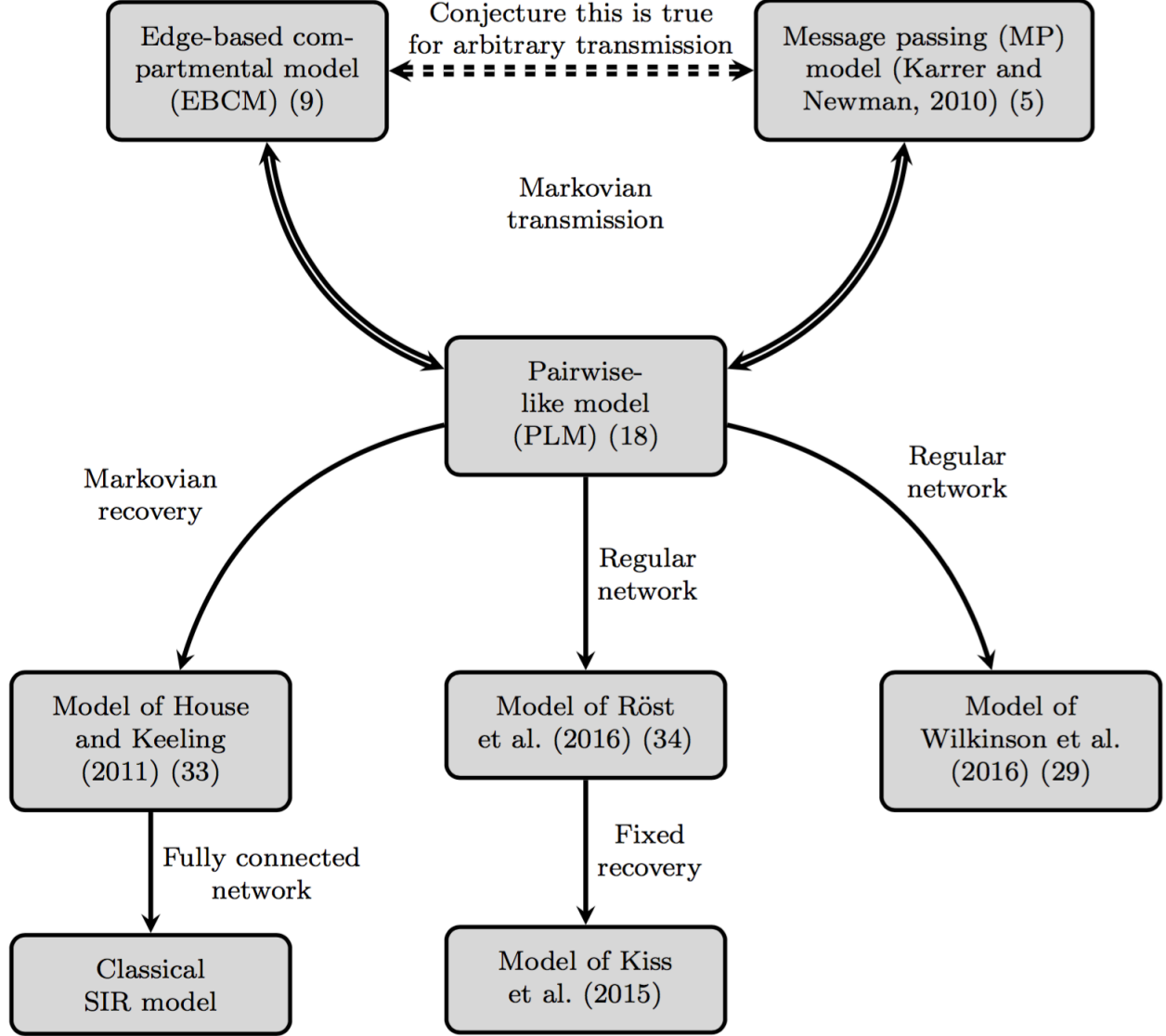


Figure 2: Diagram showing the relationship between the various models discussed in the paper. Under the assumption of Markovian/Poisson transmission, the PLM can be derived from the MP model as shown in Section 3, and the PLM and EBCM are equivalent by Theorem 2. From these proven equivalences, and the equivalence of $\Theta(t)$ and $H_1(t)$ under the most general circumstances, we conjecture that the MP and EBCM are equivalent for arbitrary transmission processes. Existing models can be derived as special cases of the PLM when the given additional assumptions are enforced, these are shown in Section 5.1 and 5.2.

5.1 Degree-regular networks

For a k -regular (homogeneous) network all nodes have the same degree, i.e. $k_u = \langle k \rangle = k$, and so the generating functions from (1) simplify to

$$G_0(x) = x^k, \quad G_1(x) = x^{k-1}, \quad \text{and} \quad G_2(x) = (k-1)x^{k-2},$$

we also introduce two new variables

$$\begin{aligned} [S](t) &= zNG_0(H_1(t)) = zN[H_1(t)]^k, \\ [SS](t) &= \langle k \rangle N (zG_1(H_1(t)))^2 = kN \left(z[H_1(t)]^{k-1} \right)^2, \end{aligned} \quad (24)$$

as the expected number of susceptible individuals, and the expected number of edges connecting two susceptible nodes, respectively. $[S](t)$ follows directly from (5), $[SS](t)$ is defined by the number of edges connecting two nodes who were both initially susceptible at time $t = 0$ and have escaped infection from their $(k-1)$ other neighbours.

Now we return to the PLM (18) and the differential equation for $[I]$ (16). Substituting in the simpler generating functions yields

$$\begin{aligned} \dot{H}_1 &= -\beta \frac{[SI]}{zkNH_1^{k-1}}, \\ [\dot{I}] &= \beta[SI] - \beta \int_0^t q(a)[SI](t-a)da - q(t)N(1-z), \\ [\dot{SI}] &= -\beta[SI][SI] \frac{(k-1)H_1^{k-2}}{zkN[H_1^{k-1}]^2} - \beta[SI] \\ &\quad + z\beta[SI](k-1)H_1^{k-2} - q(t)e^{-\beta t}(1-z)zH_1^{k-1}kN \\ &\quad - z\beta \int_0^t q(a)e^{-\beta a}[SI](t-a)(k-1)[H_1(t-a)]^{k-2} \frac{[H_1(t)]^{k-1}}{[H_1(t-a)]^{k-1}} da. \end{aligned} \quad (25)$$

This can be simplified further using (24), firstly noting that

$$\frac{[SS]}{[S]} = \frac{Nkz^2H_1^{2(k-1)}}{NzH_1^k} = kzH_1^{k-2}, \quad (26)$$

and, using \dot{H}_1 we see that

$$\frac{d(H_1^{k-1})}{dt} = -\beta(k-1) \frac{[SI]}{zkNH_1^{k-1}} H_1^{k-2} = -\beta \frac{(k-1)}{k} \frac{[SI]}{[S]} H_1^{k-1}. \quad (27)$$

Solving for H_1^{k-1} in the ODE above, using separation of variables, leads to

$$H_1^{k-1}(t) = \exp \left(-\beta \int_0^t \frac{(k-1)}{k} \frac{[SI](a)}{[S](a)} da \right). \quad (28)$$

The result of this is that the system no longer requires the message H_1 ; one can calculate the time derivatives of $[S]$ and $[SS]$ from (24). Using the new relations (26) and (28), system (25) can be

rewritten to give

$$\begin{aligned}
[\dot{S}] &= -\beta[SI], \\
[\dot{I}] &= \beta[SI] - \beta \int_0^t q(a)[SI](t-a)da - q(t)N(1-z), \\
[\dot{SS}] &= -2\beta \frac{(k-1)}{k} \frac{[SS][SI]}{[S]}, \\
[\dot{SI}] &= -\beta \frac{(k-1)}{k} \frac{[SI][SI]}{[S]} - \beta[SI] + \beta \frac{(k-1)}{k} \frac{[SS][SI]}{[S]} \\
&\quad - kNq(t)e^{-\beta t}(1-z)z \exp\left(-\beta \int_0^t \frac{(k-1)}{k} \frac{[SI](a)}{[S](a)}da\right) \\
&\quad - \beta \int_0^t q(a)e^{-\beta a} \frac{(k-1)}{k} \frac{[SS](t-a)[SI](t-a)}{[S](t-a)} \mathcal{F}(t)da,
\end{aligned} \tag{29}$$

where

$$\mathcal{F}(t) = \exp\left(-\beta \int_{t-a}^t \frac{(k-1)}{k} \frac{[SI](u)}{[S](u)}du\right). \tag{30}$$

This is identical to the system proposed by Wilkinson et al. [28]. Recently, Röst et al. [29] considered the same problem, an SIR epidemic with Poisson transmission and an arbitrary distribution of the recovery time on a regular network. By constructing an age-structured system of PDEs they were able to reach a very similar more compact model. We have, therefore, shown that the PLM extends recent work to allow for an extra level of freedom by allowing general degree distributions to be modelled in a more compact model than previous methods allowed for.

5.2 Special distributions of the recovery time

As mentioned previously, a popular choice for the recovery time distribution is to assume that times are exponentially distributed, i.e. $q(a) = \gamma e^{-\gamma a}$ for $\gamma > 0$, where $1/\gamma$ is the mean duration of infection. We briefly explain how this assumption simplifies the model and leads to familiar or well-known models. When this choice for $q(a)$ is substituted into (16), we have

$$[\dot{I}] = \beta[SI] - \gamma \left[\int_0^t e^{-\gamma a} \beta[SI](t-a)da + e^{-\gamma t} N(1-z) \right]. \tag{31}$$

Note that $e^{-\gamma a}$ is the probability of an infected node not recovering before age a , and since the number of infected nodes created a time ago is $\beta[SI](t-a)$ for $a < t$ and $N(1-z)$ is the number of initially infected nodes then (31) can be rewritten as

$$[\dot{I}] = \beta[SI] - \gamma[I]. \tag{32}$$

A similar result occurs when exponentially distributed recovery times are used in (18b), in which case the extra terms in the integral describe the probability for the susceptible node of an $[SI]$ edge to have survived until age a without receiving transmission, either along this edge or another infected neighbour. Therefore, by the same logic one can replace the final two terms in (18b) with $\gamma[SI]$. This leads to a model, which, although formulated differently, is similar to models of Volz [18] and House

and Keeling [6], namely,

$$\begin{aligned}
\dot{H}_1 &= -\beta \frac{[SI]}{z\langle k \rangle N G_1(H_1)}, \\
\dot{I} &= \beta[SI] - \gamma[I], \\
\dot{[SI]} &= -\beta[SI][SI] \frac{G_2(H_1)}{z\langle k \rangle N [G_1(H_1)]^2} - (\beta + \gamma)[SI] + z\beta[SI]G_2(H_1).
\end{aligned} \tag{33}$$

Early pairwise models assumed that the contact network was regular, and that transmission and recovery times were exponentially distributed [30]. Applying these assumptions to (29) leads to recovering such early models.

We examine another special case, when the duration of infection is a fixed period of time, σ , so that $q(a) = \delta(a - \sigma)$. When this is substituted into the PLM (18), the integral terms are non-zero only at the point $a = \sigma$, since as soon as a node is infected at time t_1 , it is known that this node will recover at exactly $t_2 = t_1 + \sigma$. This means that the system of integro-differential equations simplifies to a delay differential equation model, as stated below

$$\begin{aligned}
\dot{H}_1 &= -\beta \frac{[SI]}{z\langle k \rangle N G_1(H_1)}, \\
\dot{I} &= \beta[SI] - \beta[SI](t - \sigma) - \delta(t - \sigma)N(1 - z), \\
\dot{[SI]} &= -\beta[SI][SI] \frac{G_2(H_1)}{z\langle k \rangle N [G_1(H_1)]^2} - \beta[SI] \\
&\quad + z\beta[SI]G_2(H_1) - \delta(t - \sigma)e^{-\beta t}(1 - z)zG_1(H_1)\langle k \rangle N \\
&\quad - z\beta e^{-\beta \sigma}[SI](t - \sigma)G_2(H_1(t - \sigma)) \frac{G_1(H_1(t))}{G_1(H_1(t - \sigma))}.
\end{aligned} \tag{34}$$

This model generalises the recent work of Kiss et al. [22] to heterogeneous networks, and once again the original model in that paper can be retrieved when $q(a)$ is chosen to be a delta distribution in (29) (that original model did not explicitly account for the recovery of initially infected nodes). We illustrate all the model equivalences and reductions in Fig. 2.

Finally, it is worth briefly noting that in the case of a fully connected network, corresponding to a homogeneously well-mixed population we have that $[SI] = [S][I]$ and thus the earliest models, which assumed that the population was unstructured, can be recovered.

6 Numerical simulation of the Pairwise-like Model

In order to illustrate the accuracy of the newly derived PLM (18) we compare the numerical solution of this model to results of direct stochastic network simulation. A common approach for simulating traditional Markovian models has been to use the Gillespie algorithm [31], however, as modelling started to move away from the purely Markovian models, novel stochastic simulation methods have been derived [32, 33] which provide efficient simulation algorithms that are able to generate true sample paths of the stochastic process. In this section we take advantage of the fact that transmission remains a Poisson process in order to use an algorithm similar to those described by Barrio et al. [34]. This approach is sometimes known as the rejection method and was proven to be stochastically exact by Anderson [32]. The transmission process is run as in the standard Gillespie algorithm, and when a node becomes infected, a recovery time is drawn from the distribution $q(a)$; at each time step the time of next transmission is randomly calculated. However, if an infected node is scheduled to recover

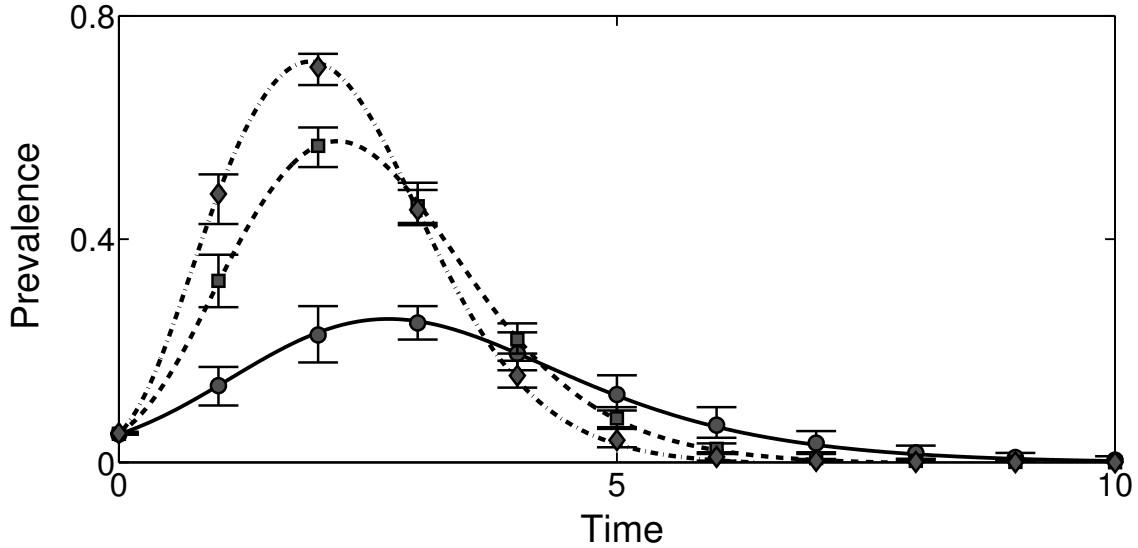


Figure 3: A plot comparing the output from the PLM (18) for an epidemic where the recovery time distribution is normal with mean 2 and standard deviation 0.5. The transmission parameter takes values $\beta = 0.15$ (solid line, circles), 0.3 (dashed line, squares) and 0.4 (dot-dashed line, diamonds), the markers represent the average of 100 simulations with corresponding parameters. The underlying networks of 1000 nodes are truncated scale-free networks with exponent 2.5 and degree bounded between 3 and 60. Error bars denote the 5th and 95th percentile.

sooner than the next planned transmission event is rejected, and the new time is updated to the next recovery time (for full details see [32]).

In the very early stages of an outbreak stochastic effects dominate the dynamics of the epidemic spread, which means that numerical simulations can often produce results that differ greatly from deterministic predictions. In this region methods such as branching process approximations [35] are more appropriate. To ensure that this does not affect our results we allow every iteration of the algorithm to reach a point where the stochastic effects are no longer a concern, and the infected population behaves deterministically. In practice this is achieved by running each individual realisation of the epidemic from a single initial seed until a specified level of infectivity is reached, at which point time is reset to zero in both the simulation and the PLM. A sufficient number of individual simulations are averaged to ensure that the mean behaviour of the stochastic model is correctly captured and is suitable for comparison with results derived from the deterministic or mean-field models. The results of such comparisons are shown for different transmission rates and various choices of the recovery time distribution in Figs. 3 (normally distributed infection times) and 4 (exponentially distributed and fixed recovery times). Firstly, in Fig. 3 the full model (18) is compared to numerical simulation on scale-free networks with normally distributed recovery times. The exceptional agreement between the simulation and the mean-field model in all cases shows strong visual evidence of this model's merits. In Fig. 4 we repeat these tests for models (33) and (34), which again show excellent agreement with numerical simulations. The agreement for networks with high degree heterogeneity and infection time distribution ranging from exponential and normal to fixed highlight the flexibility of this model.

The transmission parameters and the mean duration of infection are identical across all figures, however, the dynamics of the epidemics vary greatly in response to the change in recovery time distribution. In Fig. 4 it is clear that a fixed duration of infection causes a much larger outbreak when compared to the case of exponentially distributed recovery times, despite both of them having the same

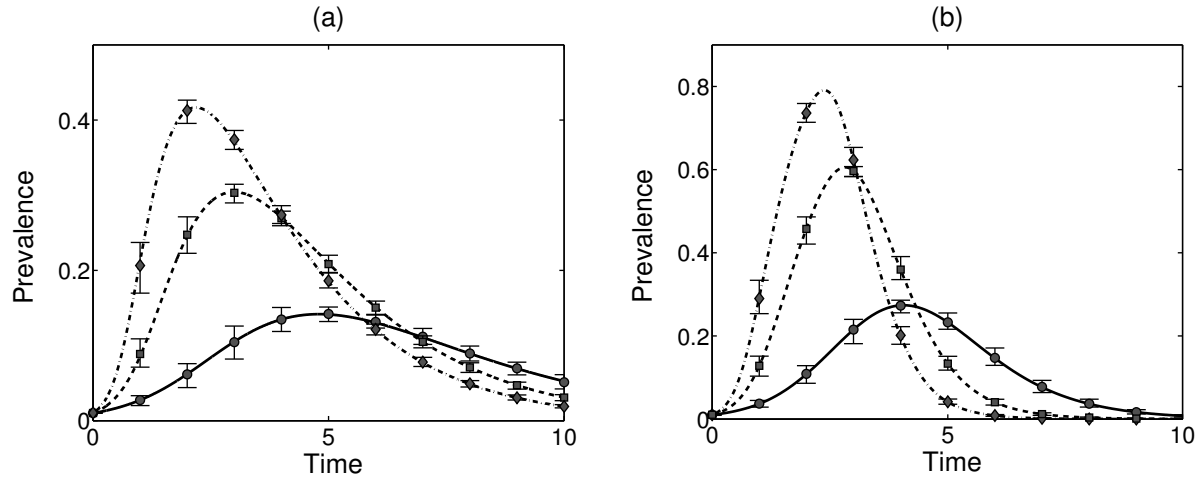


Figure 4: Tests of the models (33) and (34) on truncated scale-free networks of 10000 nodes with exponent 2.5 and degree bounded between 3 and 60. The transmission parameter takes values $\beta = 0.15$ (solid line, circles), 0.3 (dashed line, squares) and 0.4 (dot-dashed line, diamonds). In (a) results from (33) are compared to numerical simulations, with $\gamma = 0.5$ for all results. In (b) (34) is tested against numerical simulations with $\sigma = 2$. Results based on simulations are given by markers. Error bars denote the 5th and 95th percentile.

mean duration of infection. This is largely due to the change in the variance of the recovery times. In Fig. 4 (a) the exponentially distributed recovery time has the variance equal to $\gamma^{-2} = 4$, whereas, in Fig. 4 (b) the variance in recovery time is zero. This highlights the need for the true characteristics of the diseases to be accurately studied and modelled if accurate and reliable predictions are to be achieved.

7 Discussion

In this paper we have reviewed the message passing formalism for SIR epidemics on networks, and introduced a novel extension of the edge-based compartmental model to the case of arbitrary transmission and recovery processes. Both of these models are theoretically capable of accurately describing the expected dynamics of non-Markovian epidemics on tree networks; although in some cases it may be challenging to find a numerically solvable expression for the message H_1 . The final epidemic size formula found using the EBCM is in perfect agreement with results reported in existing literature [26, 7].

Adapting recent methods [25, 28] enabled us to construct a pairwise-like model (18) when the transmission process is Markovian. This compact model of only two differential equations is easier to implement numerically compared to the new EBCM and still provides an exact description of an SIR epidemic with general recovery time and degree distribution for CM networks in the limit as $N \rightarrow \infty$. Therefore, by proving that the PLM is equivalent to the EBCM we have also shown that the EBCM is equivalent to the MP model when the transmission is Markovian. We believe that this equivalence holds in a general setting, and have proved it for the message quantities, H_1 and Θ .

This new PLM model is more flexible than other pairwise models to date and thus can be applied to a wider range of data gathered from field studies. Unlike many other pairwise models, the PLM has been derived rather than heuristically defined, as is often the case [36, 37, 6]. Furthermore, the

presence of terms typically found when using the message passing formalism means that our PLM circumvents the issue of dependence on higher order arrangements, and the closure of these as found in most pairwise-like models. As a result, epidemics on networks which have a highly heterogeneous degree distribution can often lead to large systems which require many differential equations to be solved [36], or complex moment closures [38]. However, obtaining estimates on the accuracy of such approximations remains a challenge [39]. Essentially, the new PLM is a hybrid of classical pairwise and MP models.

The PLM has been studied under specific choices for both the node degree and recovery time distribution, and it has been shown that many recent and classical models can be extracted from it. By demonstrating this we hope to provide some intuition for how the PLM works and to illustrate that these newest models build on existing models but provide a modern twist. It is encouraging that such mean-field models remain relatively compact highlighting that the SIR epidemic can be modelled quite effectively, as long as a small number of key indicators about the network and the epidemic process are known. Our theoretical results have been supported by an excellent agreement between mean-field and the exact stochastic models showing that these models can be practically implemented and numerically solved, giving an excellent characterisation of the average stochastic behaviour.

Numerous extensions of the present work are possible. For example, the implementation of an efficient solver of the novel EBCM is still outstanding. Efficient numerical methods to solve such age-structured models exist but this was outside the scope of our study. In some sense the novel EBCM is the most complete mean-field model when one considers SIR epidemics on CM networks. This is due to the new model being able to handle arbitrary degree distributions, as well as arbitrary transmission and recovery processes. Additionally, such models could be refined to account for dynamic or adaptive contacts. Dynamic networks have already been incorporated in edge-based modelling in the purely Markovian setting [9], and it may be possible to extend this to a more general framework to allow for a much more unified treatment of models that include the concurrent spread of the diseases and link turnover.

Acknowledgements

N. Sherborne acknowledges funding for his PhD studies from the EPSRC (Engineering and Physical Sciences Research Council), EP/M506667/1, and the University of Sussex.

References

- [1] William O Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A*, 115(772):700–721, 1927.
- [2] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, 86(14):3200–3203, 2001.
- [3] Duncan J Watts and Steven H Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684):440–442, 1998.
- [4] Erik M Volz, Joel C Miller, Alison Galvani, and Lauren Ancel Meyers. Effects of heterogeneous and clustered contact patterns on infectious disease dynamics. *PLoS Comput. Biol.*, 7(6):e1002042, 2011.
- [5] Matthew J Keeling. The effects of local spatial structure on epidemiological invasions. *Proc. R. Soc. B.*, 266(1421):859–867, 1999.

- [6] Thomas House and Matt J Keeling. Insights from unifying modern approximations to infections on networks. *J. R. Soc. Interface*, 8(54):67–73, 2011.
- [7] Brian Karrer and Mark EJ Newman. Message passing approach for general epidemic models. *Phys. Rev. E*, 82(1):016101, 2010.
- [8] Jennifer Lindquist, Junling Ma, P Van den Driessche, and Frederick H Willeboordse. Effective degree network disease models. *J. Math. Biol.*, 62(2):143–164, 2011.
- [9] Joel C Miller, Anja C Slim, and Erik M Volz. Edge-based compartmental modelling for infectious disease spread. *J.R. Soc. Interface*, 9(70):890–906, 2012.
- [10] Leon Danon, Ashley P Ford, Thomas House, Chris P Jewell, Matt J Keeling, Gareth O Roberts, Joshua V Ross, and Matthew C Vernon. Networks and the epidemiology of infectious disease. *Interdisciplinary perspectives on infectious diseases*, 2011.
- [11] Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. Epidemic processes in complex networks. *Rev. Mod. Phys.*, 87(3):925, 2015.
- [12] Istvan Z Kiss, Joel C Miller, and Péter L Simon. *Mathematics of epidemics on networks: from exact to approximate models*. Forthcoming in Springer TAM series, 2016.
- [13] Timothy J Taylor and Istvan Z Kiss. Interdependency and hierarchy of exact and approximate epidemic models on networks. *J. Math. Biol.*, 69(1):183–211, 2014.
- [14] Joel C Miller and Istvan Z Kiss. Epidemic spread in networks: Existing methods and current challenges. *Mathematical modelling of natural phenomena*, 9(2):4–42, 2014.
- [15] Edward A Bender and E Rodney Canfield. The asymptotic number of labeled graphs with given degree sequences. *J. Comb. Theory A*, 24(3):296–307, 1978.
- [16] Mark EJ Newman. Spread of epidemic disease on networks. *Phys. Rev. E*, 66(1):016128, 2002.
- [17] Matt J Keeling and Ken TD Eames. Networks and epidemic models. *J. R. Soc. Interface*, 2(4):295–307, 2005.
- [18] Erik Volz. SIR dynamics in random networks with heterogeneous connectivity. *J. Math. Biol.*, 56(3):293–310, 2008.
- [19] NORMAN T. J. Bailey. A statistical method of estimating the periods of incubation and infection of an infectious disease. *Nature*, 174(4420):139–140, 1954.
- [20] KJ Gough. The estimation of latent and infectious periods. *Biometrika*, 64(3):559–565, 1977.
- [21] Helen J Wearing, Pejman Rohani, and Matt J Keeling. Appropriate models for the management of infectious diseases. *PLoS Med.*, 2(7):e174, 2005.
- [22] Istvan Z. Kiss, Gergely Röst, and Zsolt Vizi. Generalization of pairwise models to non-Markovian epidemics on networks. *Phys. Rev. Lett.*, 115:078701, 2015.
- [23] Neil Sherborne, Konstantin B. Blyuss, and Istvan Z. Kiss. Dynamics of multi-stage infections on networks. *Bull. Math. Biol.*, 77(10):1909–1933, 2015.

- [24] Neil Sherborne, Konstantin B. Blyuss, and Istvan Z. Kiss. Compact pairwise models for epidemics with multiple infectious stages on degree heterogeneous and clustered networks. *J. Theor. Biol.*, 407:387–400, 2016.
- [25] Robert R Wilkinson and Kieran J Sharkey. Message passing and moment closure for susceptible-infected-recovered epidemics on finite networks. *Phys. Rev. E*, 89(2):022808, 2014.
- [26] Eben Kenah and James M Robins. Second look at the spread of epidemics on networks. *Phys. Rev. E*, 76(3):036113, 2007.
- [27] Joel C. Miller. Epidemic size and probability in populations with heterogeneous infectivity and susceptibility. *Phys. Rev. E*, 76:010101, 2007.
- [28] Robert R Wilkinson, Frank G Ball, and Kieran J Sharkey. The relationships between message passing, pairwise, Kermack-McKendrick and stochastic SIR epidemic models. *arXiv preprint arXiv:1605.03555*, 2016.
- [29] G Röst, Z Vizi, and I. Z Kiss. Pairwise approximation for SIR type network epidemics with non-Markovian recovery. *arXiv preprint arXiv:1605.02933*, 2016.
- [30] Matthew J Keeling and BT Grenfell. Disease extinction and community size: modeling the persistence of measles. *Science*, 275(5296):65–67, 1997.
- [31] Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81(25):2340–2361, 1977.
- [32] David F Anderson. A modified next reaction method for simulating chemical systems with time dependent propensities and delays. *J. Chem. Phys.*, 127(21):214107, 2007.
- [33] Marian Boguná, Luis F Lafuerza, Raúl Toral, and M Ángeles Serrano. Simulating non-Markovian stochastic processes. *Phys. Rev. E*, 90(4):042108, 2014.
- [34] Manuel Barrio, Kevin Burrage, André Leier, and Tianhai Tian. Oscillatory regulation of Hes1: discrete stochastic delay modelling and simulation. *PLoS Comput. Biol.*, 2(9):e117, 2006.
- [35] JAP Heesterbeek. *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*, volume 5. John Wiley & Sons, 2000.
- [36] Ken TD Eames and Matt J Keeling. Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *Proc. Nat. Acad. Sci.*, 99(20):13330–13335, 2002.
- [37] Thilo Gross, Carlos J Dommar D’Lima, and Bernd Blasius. Epidemic dynamics on an adaptive network. *Phys. Rev. Lett.*, 96(20):208701, 2006.
- [38] Péter L. Simon and Istvan Z. Kiss. Super compact pairwise model for SIS epidemic on heterogeneous networks. *J. Comp. Net.*, 4(2):187–200, 2016.
- [39] Lorenzo Pellis, Thomas House, and Matt J Keeling. Exact and approximate moment closures for non-Markovian network epidemics. *J. Theor. Biol.*, 382:160–177, 2015.